

# Arquivos Digitais na 3ª geração da Web: uma ideia com mais de 20 anos

José Carlos Ramalho

[jcr@di.uminho.pt](mailto:jcr@di.uminho.pt)

[jcr@keep.pt](mailto:jcr@keep.pt)

VII Encontro de CTDI: Informação, Ciência e Tecnologia – fusão para a inovação



# A nossa história

Período	
1989 - 1991	HiTeX – Transcrição
1998 - 2000	Ed. Eletrónica: “Memórias de José Inácio Gavetas do Cabido”; “Bulário Bracarense”; etc.
2000	Reverse Engineering da BD das “Inquirições de Génere”
2003 – 2013	Digitarq: Gestão de metainformação, ODs, etc.
2006 – 2008	RODA: Preservação
2006 – 2013	CRAV: Consulta
2010 – 2014	SCAPE: projeto
2013 - ...	Archeevo: a nova
2013 – 2015	4C: projeto FP7

Um denominador comum:  
geração automática de índices  
(cronológico, toponímico,  
antroponímico)

Facilidades: navegação  
relacional e não hierárquica.

Produção  
massiva de  
conteúdos

Milhões de registos  
organizados  
hierarquicamente:  
ISAD, EAD, METS, etc.

# De regresso ao passado...

---

- HITEX (1989-1991): Um Sistema em Desenvolvimento para Historiadores e Arquivistas
  - Normalização: cada um deixar de fazer à sua maneira;
  - Reutilização: noção de componente reutilizável com interesse histórico;
  - Classificação: taxonomia padrão de conceitos históricos (classes) que exprime a ordem de subsunção sobre o conhecimento histórico;
  - Tolerância para com informação incompleta: permite a aquisição incremental de conhecimento histórico;
  - Resultados:
    - formato HiTeX: uma linguagem de fácil utilização para transcrição documental;
    - criação automática de índices: cronológico, toponímico e antroponímico.

# Um pequeno exemplo

“Certidão da doação que o arcebispo de Braga D.Martinho de Oliveira fez ao Cabido de Braga [: : : ] Ano de 1300.”

Início do primeiro volume do Índice das Gavetas do Cabido de Braga

Asserções:

- D. Martinho de Oliveira era o Arcebispo de Braga em 1300;
- Este arcebispo doou qualquer coisa ao Cabido de Braga, nesse mesmo ano;
- O Cabido guardou uma certidão dessa doação;
- Essa certidão está arquivada no fundo documental designado Gavetas do Cabido;
- Uma referência a essa certidão é a que se pode encontrar no fol.1, vol.1 do correspondente Índice, compilado no século XVIII.

D. Martinho de Oliveira

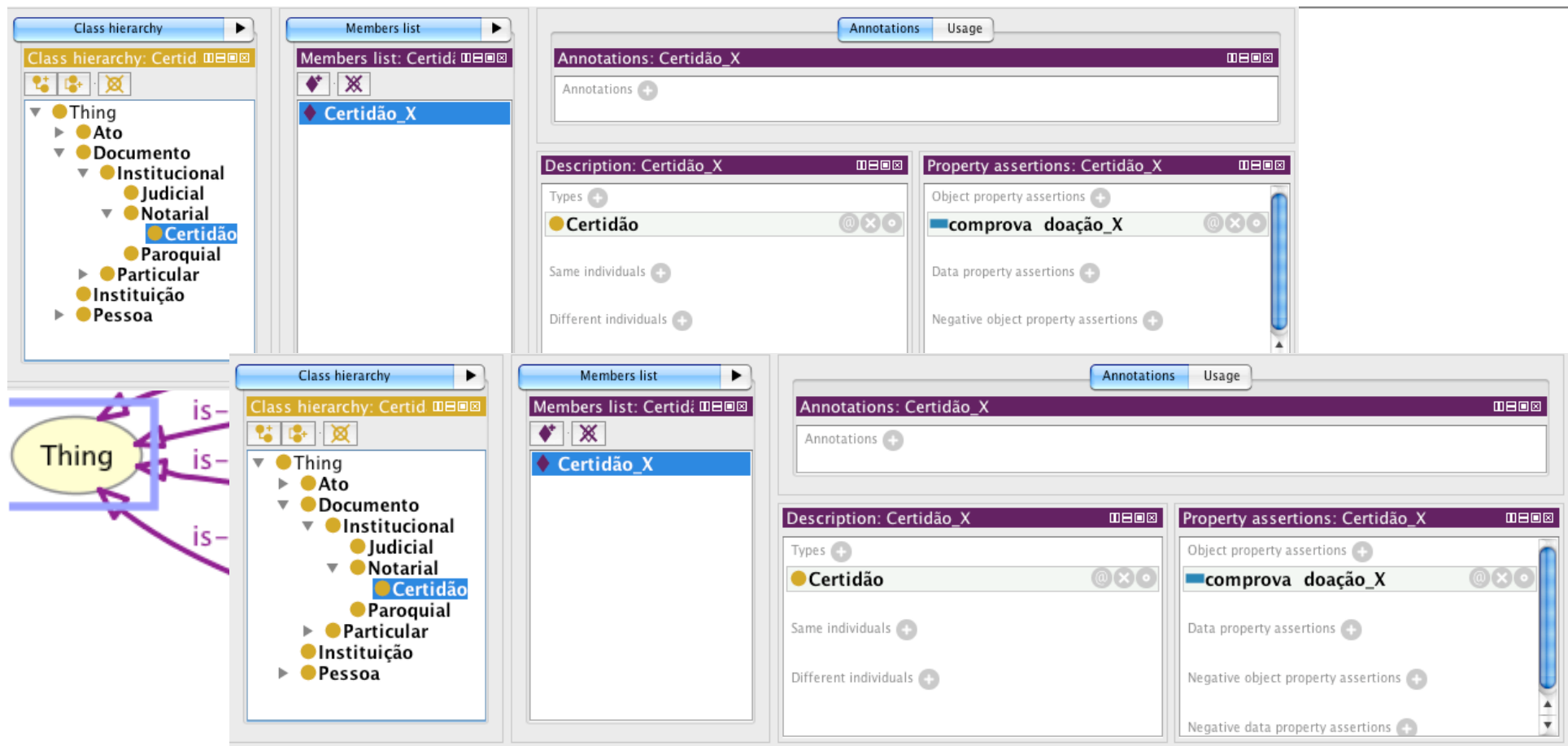
Gavetas do Cabido

Certidão X

Índice



# Modelo ontológico



# A reter...



ARQUIVO NACIONAL  
TORRE DO TOMBO



PESQUISA SIMPLES

PESQUISA AVANÇADA

CATÁLOGO

**f** TSO Tribunal do Santo Ofício 1536/1821

**sf** CG Conselho Geral do Santo Ofício 1569/1821

**sf** IC Inquisição de Coimbra

**sf** IE Inquisição de Évora 1536/1821

**sf** IG Inquisição de Goa

**sf** IL Inquisição de Lisboa 1536/1821

**sf** ILM Inquisição de Lamego 1541/1546

**sf** IP Inquisição do Porto 1536/1545

**sf** IT Inquisição de Tomar 1541/1544

**ui** 0001 Maço 1 1773/1793

**ui** 0002 Maço 2

**ui** 0003 Maço 3 1617/1756

**ui** 0004 Maço 4 1771/1793

**ui** 0005 Maço 5

**ui** 0006 Maço 6 1582/1819

**ui** 0007 Maço 7 1572/1817

**ui** 0008 Maço 8 1598/1801

**ui** 0009 Maço 9 1576/1815

**ui** 0010 Maço 10 1591/1818

**ui** 0011 Maço 11 1564/1820

**ui** 0012 Maço 12 1542/1819

## TRIBUNAL DO SANTO OFÍCIO

### NÍVEL DE DESCRIÇÃO

**f** Fundo

### CÓDIGO DE REFERÊNCIA

PT/TT/TSO

### DATAS DE PRODUÇÃO

1536 ✓ a 1821 ✓

### DIMENSÃO E SUPORTE

83.294 u.i. (3.004 liv., 79.337 proc., 624 cx., 329 mç.); perg., papel

### ENTIDADE DETENTORA

Arquivo Nacional Torre do Tombo

### PRODUTOR

Tribunal do Santo Ofício

# A Fusão dos 2 modelos

---

- Permite criar navegações alternativas à navegação hierárquica ditada pela organização dos arquivos;
- Permite criar índices automaticamente: cronológico, toponímico e antroponímico;
- Permite relacionar itens relacionamente.

# Exemplo: índice cronológico

<a href="#">Bulário Bracarense</a> <a href="#">Créditos</a> <a href="#">I - Introdução</a> <a href="#">II - Quadro Sincrónico</a> <a href="#">III - Bibliografia e Siglas</a>  <a href="#">Bulário Bracarense</a> <a href="#">Créditos</a> <a href="#">I - Introdução</a> <a href="#">II - Quadro Sincrónico</a> <a href="#">III - Bibliografia e Siglas</a>  <a href="#">1 - Referências Impressas</a> <a href="#">2 - Referências Manuscritas</a> <a href="#">3 - Abreviaturas e Siglas</a>  <a href="#">IV - Sumários</a> <a href="#">Século XI</a>  <a href="#">1</a>  <a href="#">Século XII</a>  <a href="#">2</a> <a href="#">3</a> <a href="#">4</a> <a href="#">5</a> <a href="#">6</a> <a href="#">7</a> <a href="#">8</a> <a href="#">9</a> <a href="#">10</a> <a href="#">11</a> <a href="#">12</a> <a href="#">13</a> <a href="#">14</a> <a href="#">15</a> <a href="#">16</a> <a href="#">17</a> <a href="#">18</a> <a href="#">19</a> <a href="#">20</a> <a href="#">21</a> <a href="#">22</a> <a href="#">23</a> <a href="#">24</a> <a href="#">25</a> <a href="#">26</a> <a href="#">27</a> <a href="#">28</a> <a href="#">29</a> <a href="#">30</a> <a href="#">31</a> <a href="#">32</a> <a href="#">33</a> <a href="#">34</a> <a href="#">35</a> <a href="#">36</a> <a href="#">37</a> <a href="#">38</a> <a href="#">39</a> <a href="#">40</a> <a href="#">41</a> <a href="#">42</a> <a href="#">43</a> <a href="#">44</a> <a href="#">45</a> <a href="#">46</a> <a href="#">47</a> <a href="#">48</a> <a href="#">49</a> <a href="#">50</a> <a href="#">51</a> <a href="#">52</a> <a href="#">53</a> <a href="#">54</a> <a href="#">55</a>  <a href="#">Século XIII</a>  <a href="#">56</a> <a href="#">57</a> <a href="#">58</a> <a href="#">59</a> <a href="#">60</a> <a href="#">61</a> <a href="#">62</a> <a href="#">63</a> <a href="#">64</a> <a href="#">65</a> <a href="#">66</a> <a href="#">67</a> <a href="#">68</a> <a href="#">69</a> <a href="#">70</a> <a href="#">71</a> <a href="#">72</a> <a href="#">73</a> <a href="#">74</a> <a href="#">75</a> <a href="#">76</a> <a href="#">77</a> <a href="#">78</a> <a href="#">79</a> <a href="#">80</a> <a href="#">81</a> <a href="#">82</a> <a href="#">83</a> <a href="#">84</a> <a href="#">85</a> <a href="#">86</a> <a href="#">87</a> <a href="#">88</a> <a href="#">89</a> <a href="#">90</a> <a href="#">91</a> <a href="#">92</a> <a href="#">93</a> <a href="#">94</a> <a href="#">95</a> <a href="#">96</a> <a href="#">97</a> <a href="#">98</a> <a href="#">99</a> <a href="#">100</a> <a href="#">101</a> <a href="#">102</a> <a href="#">103</a> <a href="#">104</a> <a href="#">105</a> <a href="#">106</a> <a href="#">107</a> <a href="#">108</a> <a href="#">109</a> <a href="#">110</a> <a href="#">111</a> <a href="#">112</a> <a href="#">113</a> <a href="#">114</a> <a href="#">115</a> <a href="#">116</a> <a href="#">117</a> <a href="#">118</a> <a href="#">119</a> <a href="#">120</a> <a href="#">121</a> <a href="#">122</a> <a href="#">123</a> <a href="#">124</a> <a href="#">125</a> <a href="#">126</a> <a href="#">127</a> <a href="#">128</a> <a href="#">129</a> <a href="#">130</a> <a href="#">131</a> <a href="#">132</a> <a href="#">133</a> <a href="#">134</a> <a href="#">135</a> <a href="#">136</a> <a href="#">137</a> <a href="#">138</a> <a href="#">139</a> <a href="#">140</a> <a href="#">141</a> <a href="#">142</a> <a href="#">143</a> <a href="#">144</a> <a href="#">145</a>	<p>MARIA DA ASSUNÇÃO JÁCOME DE VASCONCELOS</p> <p>Técnica Superior do ADB/UM</p> <p>ANTÓNIO DE SOUSA ARAÚJO</p> <p>Bolseiro da Fundação Calouste Gulbenkian</p>
<a href="#">127</a>	<p>[1254], MARÇO, 13, Latrão</p> <p>INOCÊNCIO IV ao Arcebispo de Bracara, ao Chantre e Mestre-Escola João de Vila Verde (Itilaviridi) cónego em Lisboa. --- BULA Sua nobis em que lhes manda tomar conhecimento e resolverasdívidas que existiam entre o Arcebispo de Braga e os Abades e conventos dos Mosteiros de Tibães, Landim, Ozo e de outros, e os reitores e clérigos da diocese que se recusavam a pagar a terça parte dos mortuários dos defuntos que enterravam nas suas igrejas e cemitérios.</p> <p>"Dat. Laterani III idus Mardi Pontificatus nostri Anno XI"</p> <p>A.D.B. --- Gav. das ReWgz5es e Mosteiros, doc. 58, orig. sem selo.</p> <p>REF.: Ferreira, 1930:38</p>
<a href="#">128</a>	<p>[1254], MAIO, 10, Assis</p> <p>INOCÊNCIO IV ao Arcebispo de Braga, D. João Egas. --- BULA Lecta coram nobis na qual, ouvida a queixa dos reitores das igrejas da diocese de Braga, mandava ao Arcebispo que proibisse aos religiosos, de qualquer Ordem a admissão aos ofícios divinos e à administração dos sacramentos, a fregueses de igrejas estranhas sem licença dos seus párocos; e não consentisse que os referidos religiosos sepultassem nas suas igrejas paro quianos doutras sem se pagarem os direitos devidos aos próprios párocos, segundo o costume, e disposições canónicas.</p> <p>"Dat. Assisii VI idus Maji Pontificatus nostri Anno XI"</p> <p>A.D.B. --- Gav. das Religi5es e Mosteiros, doc. 8, orig. com selo.</p>
<a href="#">129</a>	<p>[1254], MAIO, 11, Assis</p> <p>INOCÊNCIO IV ao Deão de Zamora. --- BULA Sua nobis na qual lhe determina, vista a queixa apresentada pelo Arcebispo de Braga, D. João Egas, que proceda com san ções canónicas contra os indivíduos que, exercendo o direito de padroado nas igrejas da província Bracarense, cometem violências e causam prejuízos aos Reitores e clérigos das suas igrejas.</p> <p>"Dat. Assisii V idus Maji Pontificatus nostri Anno XI"</p>

# Exemplo: índice toponímico

[711 712](#)

[Apêndice](#)

[ÍNDICES](#)

[Índice Antroponímico](#)

[Á](#) [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#)

[Índice Toponímico](#)

RATES --- S. Pedro [fr., c. Póvoa de Varzim, d. Porto]: *mosteiro*, doc. [57](#); *priorado*, doc. [153](#); *priores*, doc. [60](#).

REATE --- docs. [77](#), [78](#), [79](#), [80](#), [81](#), [82](#), [180](#), [181](#), [192](#).

REFOIOS DO LIMA --- Sta. Maria [fr., c. Ponto de Lima, d. Viana do Castelo]: *adabe do mosteiro*, doc. [365](#).

REMIS --- [cid. França]: doc. [18](#).

REQUIÃO --- S. Silvestre [fr., c. Vila Nova de Famalicão, d. Braga]: *prior do mosteiro*, docs. [290](#), [291](#), [292](#).

**153** [1267], JUNHO, 22, Viterbo

CLEMENTE IV ao Deão, Chantre e Mestre Rodrigo, cónego de Salamanca. --- BULA na qual lhes ordena que, indaguem das razões por que o Prior e monges do Priorado de Rates da ordem de Cluni na Diocese de Braga, não observaram o interdito que o Arcebispo de Braga D. Martinho Geraldês, pusera à sua igreja, e depois de averiguar os factos, os julguem e façam observar com firmeza as suas decisões por meio de censuras eclesiásticas.

"Dat. Viterbi Kls Juili Pontificatus nostri Anno 111"

A.D.B. --- Gav. Religiões e Mosteiros, doc. 12, orig. com selo.

**154** [1268], JULHO, 31, Viterbo

CLEMENTE IV a Mestre Guilherme Folquini, seu capelão e cónego de Narbona. --- BULA Ex parte carissirni na qual lhe refere que o Rei de Portugal D. Afonso III pediu à Santa Sé, pelos seus embaixadores, que levantasse o interdito posto pelo Arcebispo de Braga, D. Martinho Geraldês, e outros Prelados do Reino nas suas dioceses. Em virtude de tal pedido, resolveu mandá-lo como Núncio a Portugal, com autoridade para levantar o interdito, durante seis meses, a fim de resolver as questões existentes e aconselhar o Rei a evitar os vexames às pessoas eclesiásticas; de contrário, estenderia o interdito a todo o Reino.

"Dat. Viterbi lia Kls Augusti Pontificatus nostri anno quarto"

A.D.B. --- L.C., séc. XIV, fl. 33v., doc. 493 Cx. das Bulas, n. 1, doc. 14, c6p. séc. XIII.

# Exemplo: índice antroponímico

ABRANCHES, Santos --- docs. [3](#), [48](#), [49](#), [68](#), [72](#), [73](#), [74](#), [75](#), [82](#), [95](#), [97](#), [103](#), [137](#), [155](#), [156](#), [158](#), [172](#), [176](#), [177](#), [179](#), [322](#), [442](#), [639](#).

ADRIANO IV --- *papa*, doc. [23](#).

AFONSO --- docs. [73](#).

AFONSO --- *bispo de Orense*, doc. [45](#).

AFONSO --- *duque de Bragança*, docs. [333](#), [335](#).

AFONSO II --- *rei de Portugal*, docs. [67](#), [68](#), [69](#), [71](#), [72](#), [73](#), [74](#), [75](#), [158](#), [161](#).

AFONSO III --- *rei de Portugal*, docs. [111](#), [136](#), [147](#), [154](#), [155](#), [157](#), [158](#), [159](#), [163](#), [164](#), [166](#).

AFONSO IV --- *rei de Portugal*, docs. [223](#), [244](#), [251](#).

AFONSO MENDES --- *vigário*, doc. [83](#).

AFONSO V --- *rei de Portugal*, docs. [335](#), [338](#), [340](#), [341](#), [349](#), [354](#), [359](#).

AFONSO, João --- docs. [201](#), [212](#), [214](#), [227](#), [421](#), [439](#), [459](#), [485](#), [486](#), [487](#), [488](#), [489](#), [490](#), [491](#), [492](#), [493](#), [494](#), [495](#), [496](#).

**333 1458, JANEIRO, 1, S. Pedro de Roma**

*CALISTO III a D. Afonso, Duque de Bragança. --- BULA Magna fuimus na qual, depois de lembrar os agravos feitos pelo Duque ao Arcebispo de Braga D. Fernando da Guerra e as admoestações que lhe foram feitas pelos agravos referidos, aprova e confirma as penas e censuras contidas na Bula monitória de Nicolau V, se, dentro de determinado prazo, não restituir e fizer restituir à Mesa Arquiepiscopal os bens e direitos que ele e seus cúmplices lhe haviam tirado, e não prestar ao Arcebispo a obediência e reverência devidas.*

"Datum Rome apud Sanctum Petrum Anno incarnationis Dominice M CD Lo VII. KIs Januani Pontificatus nostri anno tercio"

A.D.B. --- Cx. das Bulas, n. 5, doc. 106, orig. sem selo; Gav. dos Arcebispos, doc. 94, orig. sem selo; Rer. Mem., vol. 3, séc. XVII, fl. 32; Cx. das Bulas, n. 5, doc. 107, transcrito no processo executorial de 11 de Março de 1458, orig; Cx. das Bulas, n. 5, doc. 108, transcrito na Bula Hodie a nobis da mesma data; Cx. das Bulas, n. 5, doc. 111, transcrito na Bula Provisionis nostrae de 3 de Janeiro de 1459.

REF.: Ferreira, 1930:271



- Normalização

- Ontologia: quem faz, quem certifica, ...

# Problemas



**RCAAP** | Repositório Científico de Acesso Aberto de Portugal

HOME PESQUISA AVANÇADA DIRECTÓRIO AJUDA

SUBSCREVER RSS FEED

Encontrados 417 documentos, a visualizar página 2 de 42

Pesquisa efectuada por **José Carlos Ramalho**  
Ordenado por **Relevância**

Ordenado por: Relevância | Ordem: Descendente | Actualizar

417 Resultados

Texto Pesquisado  
José Carlos Ramalho  
Actualizar

Refinar resultados

**Autor**

- ☐ Ramalho, José Carlos (111)
- ☐ Henriques, Pedro Rangel (48)
- ☐ Librelotto, Giovanni Rubert (37)
- ☐ Ferreira, Miguel (28)
- ☐ Simões, Alberto (16)

Ver mais opções

**Data**

- ☐ 2011 (54)
- ☐ 2010 (46)
- ☐ 2007 (45)
- ☐ 2009 (41)
- ☐ 2012 (37)

Ver mais opções

**Tipo de Documento**

- ☐ Documento de conferência (131)
- ☐ Artigo (88)
- ☐ Dissertação de Mestrado (79)
- ☐ Tese de Doutoramento (46)
- ☐ Outro (23)

**RODA - Repetição de Objectos Digitais Autênticos**  
Ramalho, José Carlos

Nesta palestra apresentou-se todo o processo de desenvolvimento do RODA (Repositório de Objectos Digitais Autênticos). Apresentaram-se as decisões tomadas e discutiram-se os porquês. ; Palestra apresentada no seminário "(R)evolução da informação pública : preservar, certificar e acessibilizar" Direcção-Geral de Arquivos, Torre do Tombo, Lisboa, 7 de Novembro 2011.

Data: 2011 | Origem: [Repositório - Universidade do Minho](#) [Mais info.](#)

**MuseCode : à procura da linearidade na música tonal**  
Ramalho, José Carlos

Hoje em dia, a utilização de software musical para apoio à composição, à produção de partituras e à análise musical é incontornável. Por outro lado, as notações musicais além de complexas, diferem muito uma das outras; é muito difícil para não dizer impossível, garantir a interoperabilidade entre diferentes aplicações de manipulação de música. Apesar de haver algumas iniciativas, ainda não emergiu um formato co...

Data: 2012 | Origem: [Repositório - Universidade do Minho](#) [Mais info.](#)

**Constraining content : specification and processing**  
Ramalho, José Carlos

SGML and XML are changing the way people think an act with documents. These standards have given rise to a new methodology and document processing model. Independence between structure and format, structural validation, document longevity, software and hardware independence, all these and much more became possible. However there is still a lack for content validation which relies beyond DTD scope. Recently, ...

Data: 2001 | Origem: [Repositório - Universidade do Minho](#) [Mais info.](#)

**A União Europeia e a Associação dos Países do Sudeste Asiático : o comércio com...**  
Ferreira, Carlos José Ramalho

O presente trabalho aborda a relação comercial entre a UE e a ASEAN. É efectuada a apresentação destes dois blocos económicos: a sua génese, os princípios que presidiram à sua fundação, a sua estrutura orgânica, bem como o seu percurso até à actualidade. Têm-se algumas considerações históricas acerca da ligação entre estes blocos, como representativos de dois continentes. A Globalização, a Soberania e o Poder...

Data: 2011 | Origem: [Repositório Aberto da Universidade Aberta](#) [Mais info.](#)

# Não há ontologias universais

---



# Não se pode prever

---

## o uso futuro de outros



# Evolução

Web		Arquivos pt	
1ª geração	Páginas estáticas	1ª geração	Informatização
2ª geração	Páginas geradas a partir de conteúdos em SI	2ª geração	Acesso Web à informação em regime de acesso livre
3ª geração	Sites que se atualizam automaticamente mediante alterações de 3ªs partes: BBC <ul style="list-style-type: none"><li>• Integração automática de SI</li><li>• Normalização</li><li>• Identificadores persistentes</li><li>• ...</li></ul>	3ª geração	<ul style="list-style-type: none"><li>• Integração</li><li>• Dados com semântica</li><li>• Novos modelos de acesso</li><li>• Aplicações de valor acrescentado</li></ul>

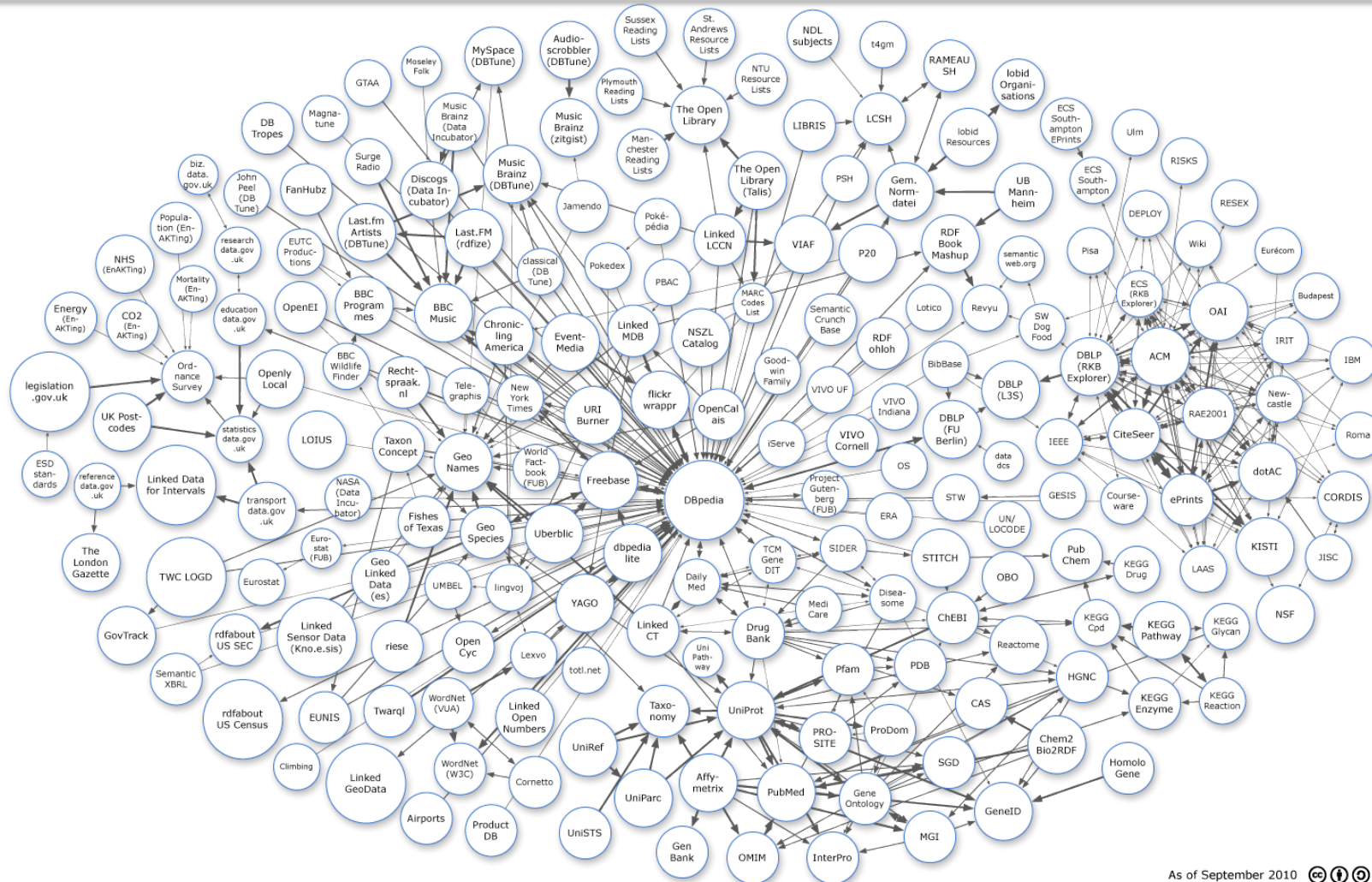
# Projecto Linked Open Data

---

- Objectivo: “expôr” repositórios de informação abertos em RDF;
- Estabelecer links RDF entre blocos de dados provenientes de diferentes repositórios;
- Criar, se possível, “SPARQL query endpoints”.



# A “cloud” LOD, Settembre 2010





# A “cloud” LOD, Setembro 2011



As of September 2011 

# Por áreas específicas



# Vocabulários

---

- Para integrar dados é necessário acordar
  - nos termos a usar:
    - “translator”, “author”
  - nas categorias a usar:
    - “Person”, “literature”
  - nas relações entre aqueles:
    - “an author is also a Person...”, “historical fiction is a narrower term than fiction”
    - novas relações podem ser deduzidas.

# O que é preciso?

---

- Acordo num vocabulário;
- Acordo numa ontologia;
- Embeber esta semântica no “front-end” dos sistemas de informação atuais (experiências em curso);
- No fim: vontade e força política (difícil mas não impossível).

# Vocabulários

---

- É necessária uma linguagem de especificação formal para estes vocabulários;
- Para os definir;
- Para lhes associar semântica de modo a tornar claro como é que novas relações podem ser deduzidas.

# Ontologia

---

„People can't **share knowledge** if they do not speak a **common language**.“ [Davenport & Prusak, 1998]

„An ontology is an **explicit specification** of a **conceptualization**.“ [Gruber, 1993]

- As Ontologias permitem melhorar a **comunicação** entre Humanos/Máquinas
- As Ontologias **normalizam** e **formalizam** o significado das palavras através de conceitos



# Ontologias (cont.)

---

- Definição:

“Especificação formal de uma área de conhecimento através da definição dos conceitos que lhe são subjacentes e relações entre estes.”

- Conhecimento explícito (fácil);
- Conhecimento implícito (difícil, é preciso torná-lo explícito).

# Para que se definem Ontologias?

---

- Para partilhar o conhecimento sobre a estrutura da informação entre pessoas ou agentes de software;
- Para tornar o conhecimento reutilizável;
- Para tornar as asserções explícitas
- Para analisar o domínio do conhecimento

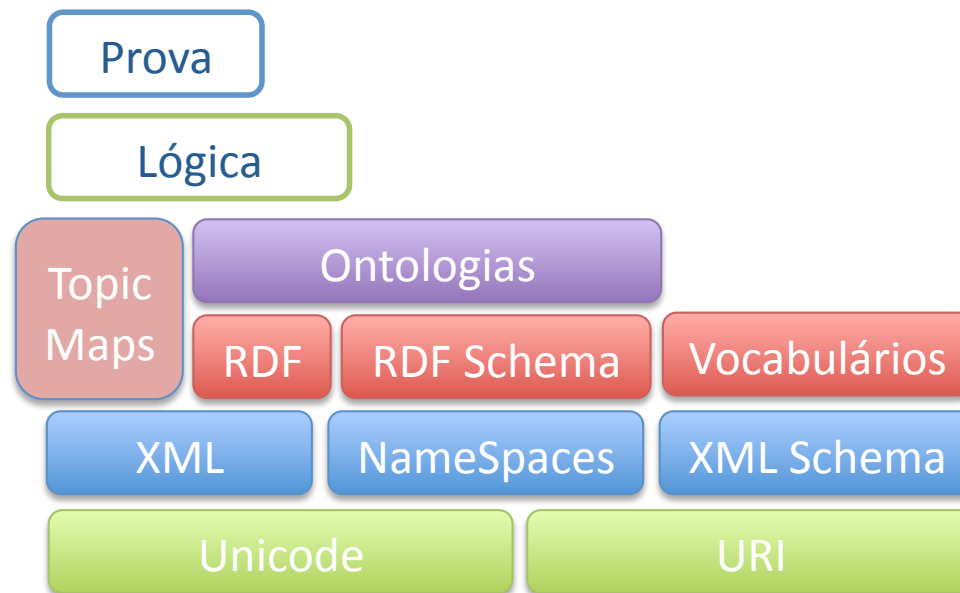
# Emergiram 3 tecnologias e 1 formato

---

- Para thesauri, glossários, etc: **SKOS**
- Para definir vocabulários mais complexos com alguma lógica subjacente: **OWL**
- E uma framework para definir regras sobre os termos e a informação: **RIF**
- **RDF/XML** para representação dos modelos semânticos

# Do bit à Semântica:

---



# Último exemplo: Inquirições de “Génere”

---

- Motor SPARQL: <http://librdf.org/query>
- Ontologia light: <http://www.di.uminho.pt/~jcr/XML/sparql/inquiricoes-light.owl>

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema>
PREFIX inq: <http://jcr.di.uminho.pt/ontologias/rc2012/inq.owl#>
```

# Interrogações: Nomes dos indivíduos?

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX inq: <http://jcr.di.uminho.pt/ontologias/rc2008#>
```

```
SELECT ?na
WHERE {
    ?a inq:nome ?na.
}
```

Count	na
1	Amaro Alvares Carvalho
2	Antonio Garcia Correia
3	Antonio Brandao Coelho
4	Antonio Alvares Carvalho
5	Agostinho Machado
6	Antonio Barbosa Veloso
7	Antonio Silva Marinho
8	Antonio Magalhaes
9	Antonio Pereira
10	Andre Vieira Carvalho
11	Agostinho Silva
12	Antonio Jorge
13	Antonio Goncalves
14	Antonio Macedo
15	Manuel Araujo Pereira
16	Alexandre Carneiro
17	Antonio Carneiro Felgueiras Gaio
18	Alexandre Barros
19	Antonio Rocha Araujo
20	Andre Coelho



# Os pares de nomes dos que são Irmãos?

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX inq: <http://jcr
```

```
SELECT ?na, ?nb
```

```
WHERE {
```

```
    ?a inq:Irmão ?b.
```

```
    ?a inq:nome ?na.
```

```
    ?b inq:nome ?nb
```

```
}
```

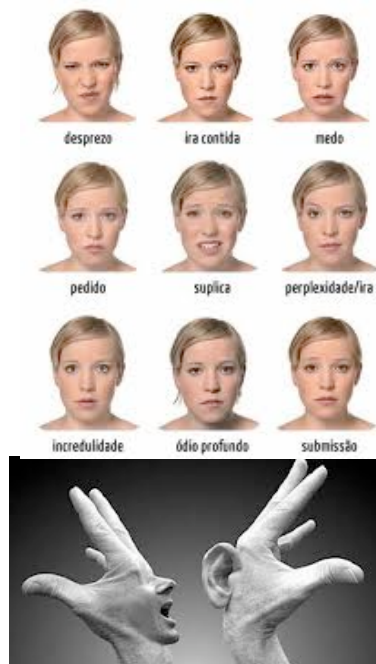
Count	na	nb
1	Joao Pereira Pinto Tavora	Joao Pereira Pinto Tavora
2	Joao Pereira Pinto Tavora	Diogo Pereira Pinto
3	Joao Pereira Pinto Tavora	Antonio Pereira Pinto Tavora
4	Diogo Pereira Pinto	Joao Pereira Pinto Tavora
5	Diogo Pereira Pinto	Diogo Pereira Pinto
6	Diogo Pereira Pinto	Antonio Pereira Pinto Tavora
7	Antonio Pereira Pinto Tavora	Joao Pereira Pinto Tavora
8	Antonio Pereira Pinto Tavora	Diogo Pereira Pinto
9	Antonio Pereira Pinto Tavora	Antonio Pereira Pinto Tavora
10	Manuel Carvalho	Manuel Carvalho

# Algumas recomendações

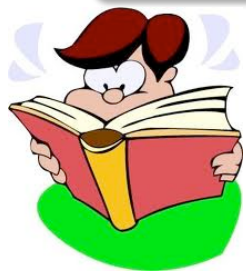
---

- Publica a tua informação primeiro, preocupa-te com as interfaces/aparência depois!
  - a “informação crua” pode ser útil assim mesmo e outros poderão usá-la;
  - podes acrescentar-lhe valor, mais tarde, criando-lhe acessos sofisticados.
- Se possível, publica a tua informação em RDF caso contrário, outros poderão ajudar-te na conversão
  - confia na comunidade...
- *Acrésceta links para outras fontes.* “Apenas” publicar não é suficiente...

# Pensamentos “soltos”



Linguagem



Informação



Estamos algures entre  
a Informação e o  
Conhecimento...



Conhecimento





---

José Carlos Ramalho

KEEPS/Universidade do Minho

[jcr@keep.pt](mailto:jcr@keep.pt) / [jcr@di.uminho.pt](mailto:jcr@di.uminho.pt)

# Questões?



<http://www.keep.pt>

# Três vias alternativas para o desenvolvimento Web

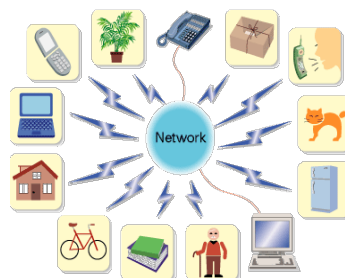


Comunidades

- Wikis
- Blogs
- Mashups
- Portais

Web 2.0

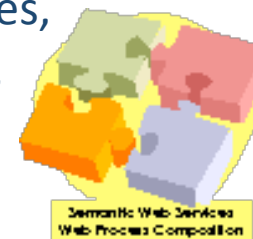
Equipamento



- Computação Ubíqua
- Espaços Inteligentes
- RFID
- Sistemas Embebidos
- Redes Sensoriais

Web of Things

Aplicações,  
Serviços,  
Agentes



- Metainformação
- Ontologias
- Serviços Web
- Agentes
- Portais Inteligentes

Semantic Web

# Informação na Web

---

- A cada dia que passa há mais informação na Web:
  - governo, saúde, conhecimento geral, informação empresarial, aviação, restaurantes,...
- Cada vez mais aplicações estão dependentes do acesso a esta informação;
- Neste momento, as aplicações estão a “ir para” a Web...



# Mas... muitas vezes, a informação está isolada em “silos”

---



# Imagina...

---

- Uma “Web” onde
  - os documentos estão acessíveis para download na Internet;
  - mas não existem hyperlinks entre eles.

# Informação disponível na Web não é suficiente...

---

- Precisamos duma infraestrutura específica para uma Web of Data real:
  - a informação está disponível na Web
    - acessível via tecnologias Web standard;
  - a informação está interligada através da Web;
  - ie, a informação pode ser integrada através da Web;
  - É aqui que aparecem as tecnologias da Web Semântica.

# Ideia: ligar os silos

---



Photo credit "kxllly", Flickr

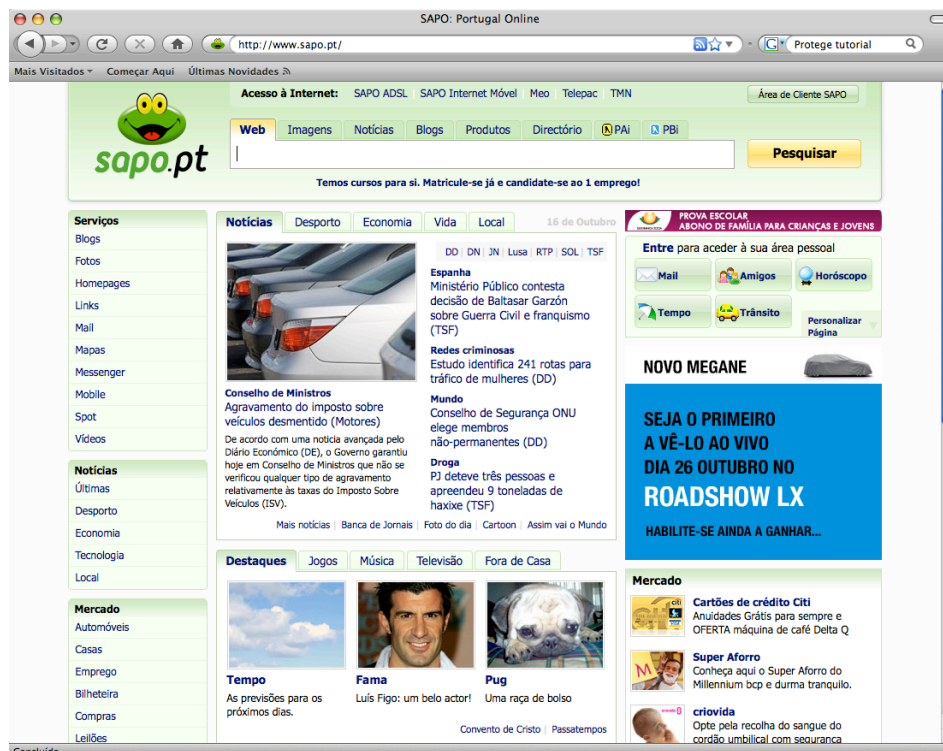
# Semantic Web

---

- A Web tornou-se possível devido a algumas normas
    - **TCP/IP** para o transporte de bits através de redes
    - **HTTP & HTML** para transportar e formatar conteúdos em formato hipertexto
  - e a aplicações capazes de explorar este cenário
  - **1ª geração:** páginas HTML escritas à mão
  - **2ª geração:** páginas geradas por máquinas
    - dirigidas ao processamento e interacção humanos
  - Na próxima geração, os recursos devem estar acessíveis a processos automáticos
-



# Qual é o problema?



Uma página Web é composta por:

- informação de formatação (e.g. tamanho de letra, cor);
- links para conteúdos relacionados.

O Conteúdo semântico é acessível aos humanos mas não aos computadores (pelo menos de um modo fácil).

# A informação que vemos...



## José Carlos Ramalho

*Prof. Auxiliar*

Department of Informatics // CCTC Research Center

University of Minho

Campus de Gualtar 4700-057 Braga - Portugal

Phone: +351 253 604479 | FAX: +351 253 604471 | Ext.: 604479

jcr @ di.uminho.pt | <http://www.di.uminho.pt/~jcr/>

[Personal information](#) | [Research projects](#) | [Publications](#) | [Communications](#) | [Thesis Supervision](#) | [Thesis Examination](#) | [Conferences](#) | [Teaching](#) | [Personal hobbies](#)

## Thesis Examination

### Phd (as Examiner)

- Francisco Javier Calzada Prado, "*Los Objetos de Aprendizaje en la Educación Superior: Fundamentos, Tratamiento y Recuperación en Repositorios y Bibliotecas Digitales Educativas*", Doctorado, Madrid, Spain, Jun. 2009.
- Jana Jagodic, "*The Role of an ICT Change Agent in ICT Diffusion within Technology Projects in Public and Private Sector Settings*", Doctorate of Business Administration (DBA), School of Business, University of Ballarat, Australia, August 2008.
- Gustavo Vasconcelos Arnold, "*Automatization of the Code Generation for Different Industrial Robots*", Escola de Engenharia da Universidade do Minho, June 2008.
- Maria João Tinoco Varanda Pereira, "*Sistematização da Animação de Programas*", Escola de Engenharia da Universidade do Minho, Oct. 2003.

### Msc (as Examiner)

- João Miguel Silva, "*Sharing Botanical Information using Geospatial Databases*", na Faculdade de Engenharia da Universidade do Porto, July 2009.
- Pedro Miguel Silva, "*Computação Ubíqua no Ensino Básico: Plataforma orientada à informação multissensorial georreferenciada*", na Faculdade de Engenharia da Universidade do Porto, July 2009.
- Marília Natália Gonçalves, "*Earthquake Record Selection for Dynamic Structural Analysis: an Application with a Web Interface*", na Faculdade de Engenharia da Universidade do Porto, July 2009.
- Luis Matias, "*Sistema de Gestão Financeira para um Parque Tecnológico utilizando um Sistema de Geração Automática de Código*", na Faculdade de Engenharia da Universidade do Porto, March 2009.



# A informação que as máquinas vêm...

---

0000040	042057	072141	037145	020012	020040	036040	060503	060564
0000060	067554	072547	051545	060564	072564	037163	071104	063141
0000100	036164	041457	072141	066141	063557	062565	072123	072141
0000120	071565	005076	020040	020040	041474	062562	072141	062145
0000140	030076	027466	027470	034461	034471	027474	071103	060545
0000160	062564	037144	020012	020040	036040	062522	067560	064563
0000200	067564	074562	040476	052116	036124	051057	070145	071557
0000220	072151	071157	037171	020012	020040	036040	064524	066164
0000240	037145	064103	067141	062543	060554	064562	020141	062544
0000260	042040	020056	060515	064562	020141	036111	052057	072151
0000300	062554	005076	020040	020040	042474	072170	067145	037164
0000320	034464	066040	073151	036056	042457	072170	067145	037164
0000340	020012	020040	036040	067515	064544	064546	062145	030076
0000360	027467	033061	031057	030060	036070	046457	062157	063151
0000400	062551	037144	020012	020040	036040	062522	047146	037157
0000420	052120	052055	026524	027501	027506	027501	036062	051057
0000440	063145	067516	005076	020040	020040	047074	072157	071545
0000460	047076	072157	020141	067541	062440	062554	062555	072156
0000500	020157	062544	064440	063156	071157	060555	123703	121703
0000520	020157	061442	131703	064544	067547	062040	020145	062562
0000540	062546	141562	067252	064543	021141	020072	141543	062263
0000560	063551	020157	062544	071040	063145	071145	125303	061556
0000600	060551	060440	072156	063551	035157	050040	026524	052124
0000620	046455	047503	041457	046517	041457	041517	031057	036056

---

# Limitações actuais da Web

---



Interacção Homem-Máquina



Interacção Máquina-Máquina

# Exemplo: pesquisa genérica

The screenshot shows a Google search interface. The search bar contains the text "José Carlos Ramalho artigos". Below the search bar, the word "Pesquisar" is visible. A blue oval highlights the search results summary: "Cerca de 6.500.000 resultados (0,24 segundos)". On the left side, there is a vertical menu with links: "Tudo", "Imagens", "Mapas", "Vídeos", "Notícias", "Compras", "Mais", "Pesquisar na Web", "Pesquisar páginas em Português", and "Páginas estrangeiras traduzidas". The main content area displays the search results for "José Carlos Ramalho". It includes a link to "di.uminho.pt/~jcr", a brief biography of José Carlos Ramalho as a professor at the University of Minho, and a link to a bibliography. Below this, there is a link to "José Carlos Ramalho - Perfil do Google" with a URL. A small profile picture of José Carlos Ramalho is shown next to a bio snippet, with an "Editar perfil" button below it.

Google

José Carlos Ramalho artigos

Pesquisar

Cerca de 6.500.000 resultados (0,24 segundos)

Tudo

Imagens

Mapas

Vídeos

Notícias

Compras

Mais

Pesquisar na Web


Pesquisar páginas em Português

Páginas estrangeiras traduzidas

**José Carlos Ramalho**  
[di.uminho.pt/~jcr](http://di.uminho.pt/~jcr)  
**José Carlos Ramalho.** Prof. Auxiliar Department of Informatics // CCTC Research Center University of Minho Campus de Gualtar 4700-057 Braga - Portugal ...  
[Communications](#) - [Publications](#) - [Thesis Examination](#) - [Thesis Supervision](#)

**Bibliography generated from jj.bib**  
[natura.di.uminho.pt/~jj/bib/jj.bib.html](http://natura.di.uminho.pt/~jj/bib/jj.bib.html)  
In **José Carlos Ramalho**, João Correia Lopes, and Luís Carriço, editors, XATA 2007 — 5ª ... Neste **artigo** propomos uma metodologia de ordem superior para o ...

**José Carlos Ramalho - Perfil do Google**  
<https://plus.google.com/114789018167262122511>

 Viana do Castelo - Engenharia de sistemas e informática ...  
"Always pushing the limits..."  
Sou professor de programação de computadores no ...

Editar perfil

# Exemplo: pesquisa no domínio

The screenshot shows the RCAAP (Repositório Científico de Acesso Aberto de Portugal) website. The header includes the RCAAP logo and navigation links: HOME, PESQUISA AVANÇADA, DIRECTÓRIO, AJUDA. A banner image shows five glass bottles with colored liquids (blue, orange, green, red, yellow) and a broken bottle. The search results section is titled 'Encontrados 266 documentos, a visualizar página 1 de 27'. It shows the search criteria: 'Pesquisa efectuada por José Carlos Ramalho' and 'Ordenado por Relevância'. The results are sorted by 'Relevância' in 'Descendente' order. A blue oval highlights the 'Actualizar' button. Below the search bar, there are two search results. The first result is 'A foundation for automatic digital preservation' by Ferreira, Miguel; Baptista, Ana Alice; Ramalho, José Carlos. The second result is 'Qualidade na publicação electrónica : como controlá-la?' by Ramalho, José Carlos; Henriques, Pedro Rangel. On the right side, there is a sidebar titled '266 Resultados' with a search bar containing 'José Carlos Ramalho' and an 'Actualizar' button. Below this, there is a 'Refinar resultados' section with a list of authors and their document counts: Ramalho, José Carlos (96), Henriques, Pedro Rangel (48), Librelotto, Giovanni Rubert (37), Ferreira, Miguel (23), and Almeida, J. J. (10). A link 'Ver mais opções' is at the bottom of the sidebar.

Contatos English

RCAAP | Repositório Científico de Acesso Aberto de Portugal

HOME PESQUISA AVANÇADA DIRECTÓRIO AJUDA

SUBSCREVER RSS FEED

Encontrados 266 documentos, a visualizar página 1 de 27

Pesquisa efectuada por **José Carlos Ramalho**  
Ordenado por **Relevância**

Ordenado por: Relevância Ordem: Descendente Actualizar

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 rss feed resultados pesquisa

**A foundation for automatic digital preservation**  
Ferreira, Miguel; Baptista, Ana Alice; Ramalho, José Carlos  
Miguel Ferreira, Ana Alice Baptista and José Carlos Ramalho propose a Service-Oriented Architecture to help cultural heritage institutions to accomplish automatic digital preservation. ; Fundação para a Ciência e a Tecnologia (FCT)  
Data: 2006 | Origem: [RepositóriUM - Universidade do Minho](#) [Mais info.](#)

**Qualidade na publicação electrónica : como controlá-la?**  
Ramalho, José Carlos; Henriques, Pedro Rangel  
Hoje em dia assiste se a uma utilização desenfreada de suportes digitais para uma publicação electrónica selvagem e ditada pelas regras de competição do mercado. Assim, é fácil encontrar CDROMs ou páginas WWW contendo vários erros, que podem ser simples erros ortográficos ou erros semânticos que podem ter consequências graves. De facto, se nos debruçarmos um pouco sobre o ciclo de vida da publicação electróni...  
Data: 1998 | Origem: [RepositóriUM - Universidade do Minho](#) [Mais info.](#)

**266 Resultados**

☐ Texto Pesquisado  
José Carlos Ramalho  
Actualizar

**Refinar resultados**

☐ Autor

- ☐ Ramalho, José Carlos (96)
- ☐ Henriques, Pedro Rangel (48)
- ☐ Librelotto, Giovanni Rubert (37)
- ☐ Ferreira, Miguel (23)
- ☐ Almeida, J. J. (10)

[Ver mais opções](#)

# Exemplo: serviços de valor acrescentado



## SCEUR workbench

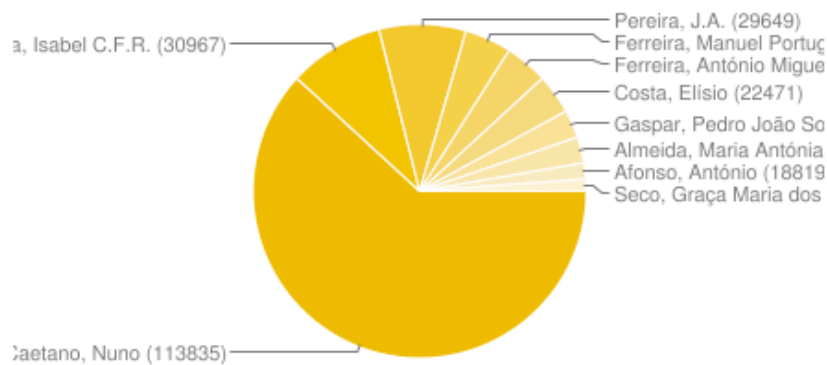


Serviço Centralizado de Estatísticas de Utilização de Repositórios

### 1) Tipo de eventos

Selecionar o tipo de evento **Gráfico gerado**

- ☐ Consultas de meta
- ☒ Downloads de Doc
- ☐ Consultas de meta
- ☐ Depósitos



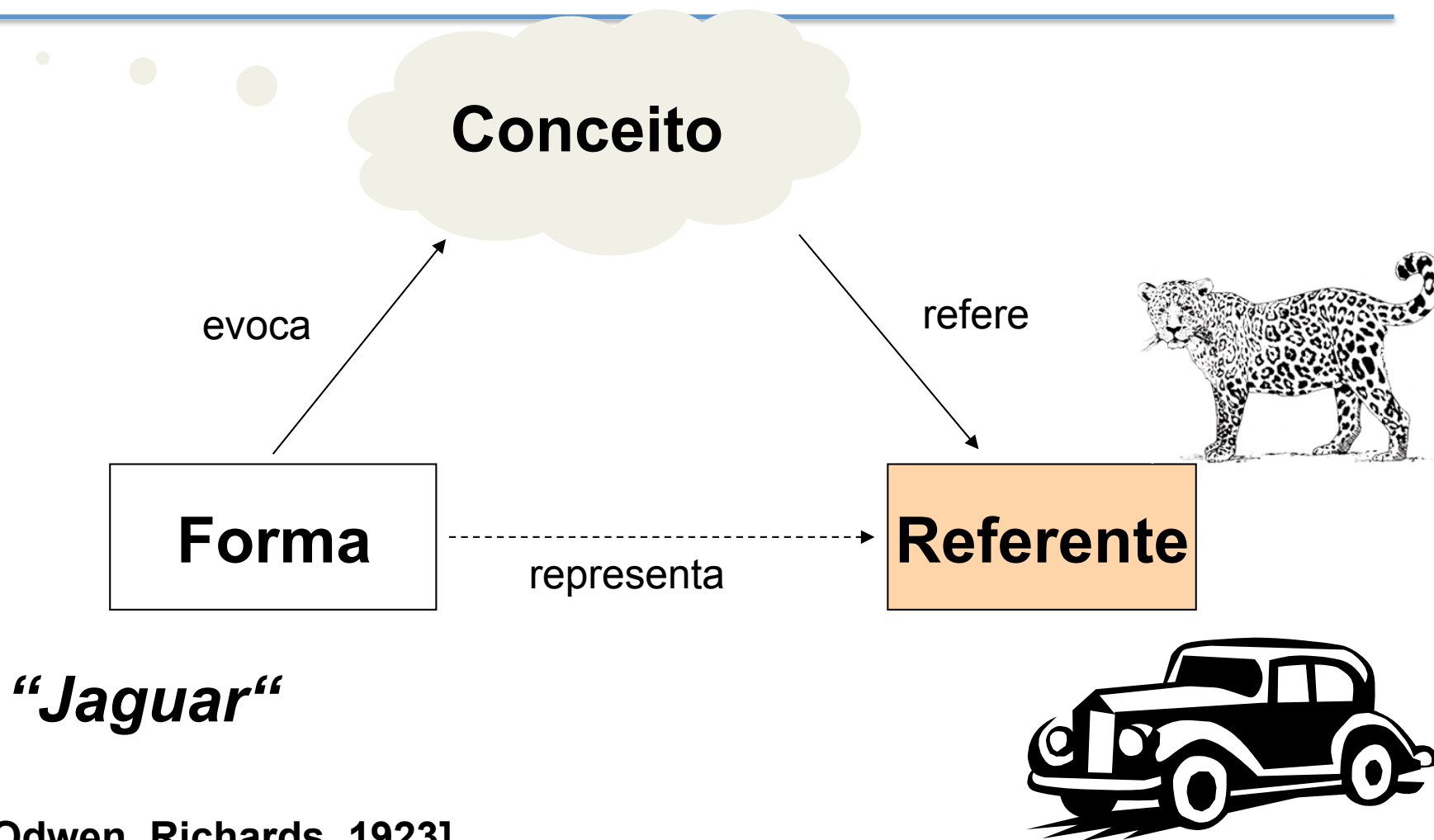
Chave	Valor
Caetano, Nuno	113835 (1.53%)
Ferreira, Isabel C.F.R.	30967 (0.42%)
Pereira, J.A.	29649 (0.40%)
Ferreira, Manuel Portugal	23512 (0.32%)
Ferreira, António Miguel Pereira Jorge	22924 (0.31%)
Costa, Elísio	22471 (0.30%)
Gaspar, Pedro João Soares	20643 (0.28%)
Almeida, Maria Antónia de Figueiredo Pires de	20361 (0.27%)
Afonso, António	18819 (0.25%)
Seco, Graça Maria dos Santos Batista	18191 (0.24%)
Total	321372 (4.31%)

# Qual é o problema?

---

- Quem é quem?
- Registo de autoridade, certificação da identidade...
- Unificação de várias formas para o mesmo referente.

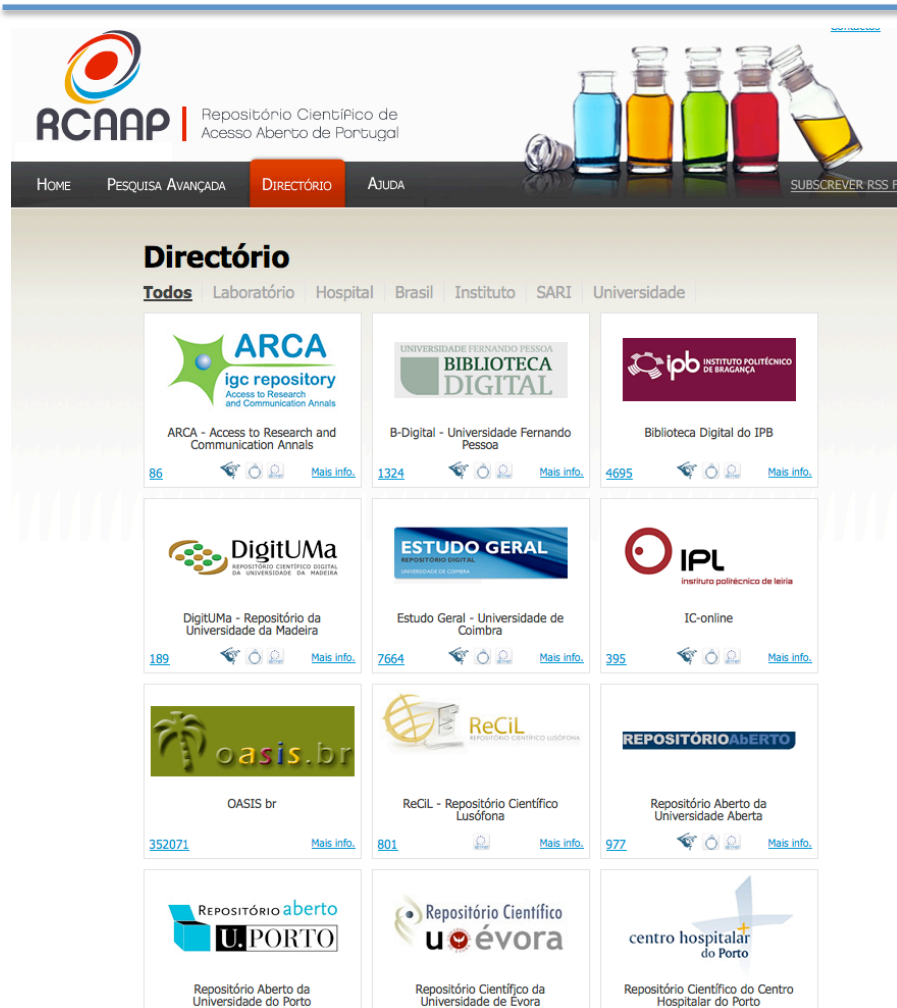
# Triângulo do Significado



[Odwen, Richards, 1923]



# Possível solução: “ir a todos”



Muitos silos!

# A abordagem da Web Semântica

Integrar as várias bases de dados...

- ARCA
- DigitUMA
- RepositoriUM
- Repositório Aberto
- ...



# Integração de Dados

---

- Mapeiam-se os dados provenientes das várias fontes de informação numa representação abstracta:
  - a informação deverá ser independente da sua representação...
- Fundem-se as representações resultantes;
- A partir daqui podem-se fazer queries nesta representação resultante!
  - queries que não seriam possíveis nas fontes isoladas.

# Então onde está a Web Semântica?

---

- A Web Semântica fornece as tecnologias que tornam possível esta integração!

# Designações

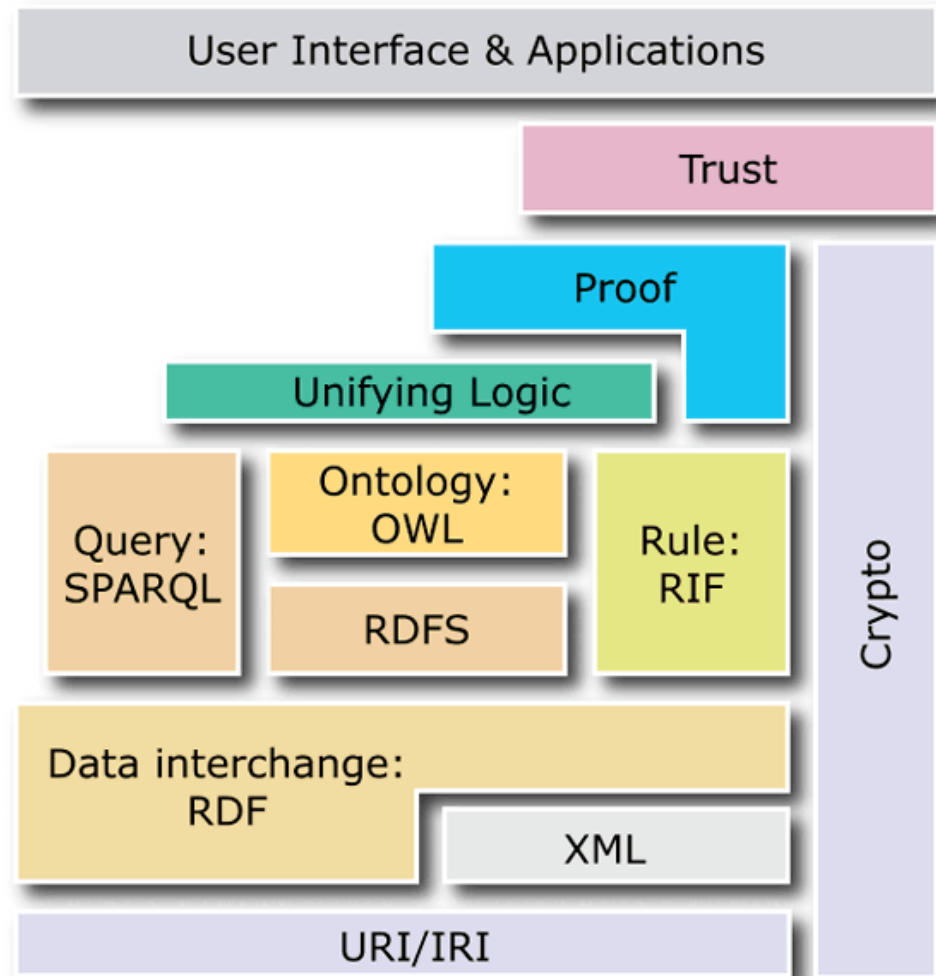
---

- Semantic Web
- Web of Data
- Giant Global Graph
- Data Web
- Web 3.0
- Linked Data Web
- Semantic Data Web

- Um denominador *comum de* tecnologias:
    - ...permite aplicações diversas
    - ...suporta a interoperabilidade
  - Um conjunto *coerente* de tecnologias:
    - ...permite a sua aplicação incremental
    - ...fornece uma base para a inovação
  - Um conjunto *normalizado* de tecnologias:
    - ...reduz a possibilidade de bloqueio por 3<sup>as</sup> partes
    - ...promove a escolha de várias ferramentas em alternativa
-

# A pilha de linguagens

- O trabalho nesta área tem-se focado na definição de uma pilha de linguagens.
- Estas linguagens são usadas para suportar a representação e o uso de metainformação





# Semantic Web: evolução temporal

